Signal Processing Compact Binary Coalescence Gravitational Wave Data from the Advanced LIGO Detectors

A Thesis Presented in Partial Fulfillment of
the Bachelor of Science in Physics
and
the Bachelor of Science in Astronomy

Sophia Elissa Bogat

Abstract

This thesis explores the complex signal processing tools and techniques used to perform gravitational wave astronomy. The first ever direct observation of spatial strain caused by a gravitational wave was achieved by the Laser Interferometer Gravitational-Wave Observatory (LIGO) on September 14th, 2015, nearly 100 years after Albert Einstein predicted their existence. Because the amplitude of the strain is so small (on the order of 10^{-21}), it must be measured by a 4 kilometer long interferometer equipped with extremely advanced thermal and seismic vibration isolation systems. Furthermore, the data must undergo significant processing in the form of whitening, matched filtering, and bandpass filtering. We present a detailed study of the steps undergone to identify and validate potential gravitational wave signals using the LIGO-designed PyCBC software framework for the observation of compact binary coalescence.

Department of Physics
Under the supervision of Dr. Richard Matzner
University of Texas at Austin
December 2017

Contents

List of Figures

Acknowledgments

I would like to offer my sincerest thanks to Dr. Richard Matzner for supervising the completion of this thesis along with my journey into the study of general relativity.

I would also like to thank Alex Buchanan, Mark Selover, and Dr. Justin Feng, who welcomed me into the relativity group and gave me every conceivable crash course. Thank you for being an invaluable help with this endeavor.

Thanks as well to the LIGO Scientific Collaboration for providing the resources, data, documentation, and inspiration to complete this project. I know there are many in the scientific community who have benefited from their openness and generosity.

To Dr. Greg Sitz, thank you for all you do for the Physics Department, and for your many wise words over the years.

Thank you, Dr. Sonia Paban, for an incredible education and for inspiring me to be a teacher as well as a scientist.

And to Kendra Michelle, Fallon Scheibel, and Isabella Palacios, thank you for your unwavering support, and for your love always. To my brother Antonio, my mother Carrie, and my step-father Dave, thank you for believing in me and always having such high expectations. And to my dad, Leni Bogat, thank you for instilling within me the passion for discovery and the drive to create that has carried me through this adventure.

Sophia Elissa Bogat
Austin, Texas
December 2017

1 Introduction

On September 14th, 2015, the Advanced Laser Interferometer Gravitational-Wave Observatory (LIGO) made the first ever detection of a gravitational wave. This signal, named GW150914, has confirmed predictions made by Albert Einstein in 1916, and its detection has given life to the new field of gravitational wave astronomy [1].

Gravitational waves are waves of transverse spatial strain, in which space itself is stretched in one direction and compressed in a perpendicular direction, as illustrated in Figure 1. They are produced by accelerating masses, in many ways analogous to electromagnetic waves produced by accelerating electric charges, however, the magnitude of the strain (the fractional change in length) caused by gravitational waves is extremely small. So small, in fact, that after Einstein predicted them in 1916, their existence was widely debated for the next 41 years [2]. The maximum strain produced by the GW150914 event was on the order of 10^{-21}, indicating that the change in the length of space was about a thousandth of a billionth of a billionth of the original length. To achieve such astounding precision, LIGO has made use of extremely advanced technological systems and complicated data processing techniques [3]. In this thesis, I will first describe the necessary and relevant parts of general relativity and its predictions, as well as the physical measurement techniques employed by LIGO. Then, I will explore LIGO's data analysis process in an attempt to demystify the final steps that allow us to use this completely new window into our universe.

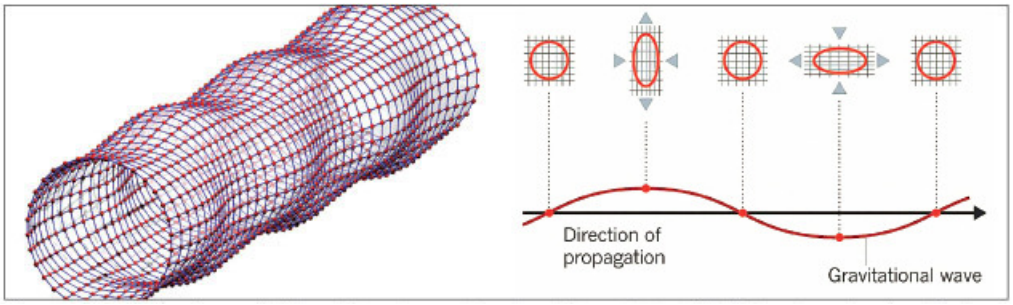


Figure 1: This image shows the effect on space-time of a passing gravitational wave (GW). If we define a cylindrical region of space, then a wave traveling along its axis would create distortion in alternating directions perpendicular to the axis and perpendicular to each other. The frequency of this alternation is determined by the source of the wave. As an example, if the source is a compact binary system, then the gravitational wave frequency is twice the orbital frequency of the binary system [4]. Image credit: Copyright ©2016 Max Planck Institute for Gravitational Physics, Golm/Potsdam .

1.1 General Relativity and Gravitational Waves

Einstein's theory of general relativity (GR) is, in essence, a description of the relationship between matter distribution and space-time curvature [5]. In order to describe each of these components completely, we use mathematical elements called tensors, and the associated mathematical branch of tensor calculus. Tensor equations have the extremely useful property that they are "coordinate invariant," that is, they hold for all coordinate systems.

It is important first to define events in space-time using four-vectors, which are vectors of the form $V = (ct, x, y, z)$. The components of four vectors may be written as V^μ, where the raised index represents an arbitrary coordinate. Conventionally, the index 0 refers to t, 1 refers to x, 2 refers to y, and 3 refers to z. We will also use units in which the speed of light $c = 1$ [5].

The tensor that describes the relationship between the coordinates is called the metric, $g_{\mu\nu}$, which is used to define the norm of a vector, the dot product of two vectors, and the infinitesimal distance between

two points (the space-time interval) [6]. For classical "flat" space, it can be written as the matrix,

$$\eta_{\mu\nu} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{1}$$

where the indices are lowered, but represent the same coordinates in the same order. In this case, the metric is called the Minkowski metric, and the dot product is defined as

$$A \cdot B \equiv \eta_{\mu\nu} A^\mu B^\nu = -A^0 B^0 + A^1 B^1 + A^2 B^2 + A^3 B^3. \tag{2}$$

Thus, the metric can be written as the space-time interval,

$$ds^2 \equiv \eta_{\mu\nu} dx^\mu dx^\nu = -dt^2 + dx^2 + dy^2 + dz^2, \tag{3}$$

where the raised 2 is an exponent, not an index [5].

These equations can be recognized as the familiar dot product and space-time interval of special relativity. When we generalize the metric to be $g_{\mu\nu}$, we can describe these characteristics for a curved space-time. The Einstein field equations for general relativity, summarized below, can then be thought of as the equations of motion for the metric [5].

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 8\pi G T_{\mu\nu} \tag{4}$$

The left-hand side of this equation carries information about the curvature of space-time, containing the metric, as well as the terms $R_{\mu\nu}$ and R, which are constructed from the metric and its derivatives. The right-hand side contains Newton's gravitational constant G, and the energy-momentum tensor $T_{\mu\nu}$, which fully describes energy and momentum of the matter fields. Thus, we can see that the distribution of matter and energy affects the curvature of space-time, and the curvature of space-time directs the movement of matter and energy [5].

Einstein published this result in 1915 [7], and in the following year, two very important conclusions were drawn. One is that Karl Schwarzschild found a unique solution to the field equations for a spherically symmetric matter distribution in a vacuum, whose metric can be written in spherical coordinates as

$$ds^2 = -(1 - \frac{2Gm}{r})dt^2 + (1 - \frac{2Gm}{r})^{-1} dr^2 + r^2(d\theta^2 + \sin\theta d\phi^2), \tag{5}$$

where we consider a radius r containing a mass m. We can see that the metric components go to infinity at $r = 0$ and $r = 2Gm$, implying an ill-defined region of space-time, called a singularity. However, a clever change of coordinates removes the problem at $r = 2Gm$, leaving $r = 0$ as the only true, physical singularity [8].

The Schwarzschild solution accurately describes any spherically symmetric body whose mass is entirely contained within r, but it does not accurately describe the interior of these objects. Thus, for bodies such as planets and stars, whose radii are far greater than $2Gm$, we do not have to worry about the presence of singularities. If we consider a body whose radius R is less than $2Gm$, the metric tells us that for any particle within the boundary $r < 2Gm$, r becomes a time-like coordinate, with all paths leading to the point $r = 0$. This means that for any body with radius $R < 2Gm$, all of its mass must collapse to the singularity at $r = 0$. Such a body is called a black hole, as the geometry of space-time within the surface $r = 2Gm$ (the "event horizon"), is such that neither mass nor radiation can escape [8].

The second significant conclusion was drawn by Einstein himself. Upon linearizing the field equations, he found a plane wave solution that described "ripples" in space-time that propagate at the speed of light [9]. Decades later, plane wave and spherical wave solutions to the fully non-linear equations were found by physicists including Hermann Bondi, Ivor Trautman, and Andrzej Robinson. These solutions were interpreted physically as gravitational radiation- transverse waves that can remove gravitational energy from a system [10, 11]

Gravitation radiation is produced by accelerating masses, in many ways analogous to electromagnetic radiation. However, gravitational radiation has no dipole component. This is because total linear momentum

is conserved, so the first derivative of the mass dipole moment is constant, and the second derivative vanishes, producing no radiation [9]. The lowest order source of gravitational waves is therefore the quadrupole, in agreement with the formula that Einstein found in his linear approximation:

$$4\pi r^2 t_4^r = \frac{32}{5} I^2 \omega^6. \tag{6}$$

Here t_4^r represents the radial energy component of the gravitational field, I represents the mass quadrupole moment tensor, ω represents the angular velocity of the body, and the left-hand side of this equation is the rate of energy loss due to the radiation. Einstein also calculated the strain caused by a quadrupole radiation source at a distance d to be

$$h = \frac{2G}{c^4 d} \frac{\partial^2 I}{\partial t^2}, \tag{7}$$

where we see Newton's gravitational constant G and the speed of light c in the constant of proportionality. Given that G is very small and c is very large, it is apparent that the magnitude of the strain must be incredibly minuscule [9]. The question then becomes: What are the most probable sources for detectable gravitational waves?

1.1.1 Compact Binary Coalescence

By inspection of equation (7), we can see that the strain increases with both the mass and acceleration of a system, so systems of high-mass objects with large acceleration would be good candidates for detectable gravitational radiation. The most promising cases are binary systems of compact objects, namely neutron stars and black holes, whose orbits will decay to coalescence due to gravitational radiation [12]. The gravitational waves produced by these systems have a frequency of twice the orbital frequency of the masses, and carry energy away from the system at a rate given by equation (6). Thus, at a distance $r \times 100$ Mpc from the system we would be able to detect an oscillating gravitational waveform with r.m.s. strain

$$\langle h \rangle \approx 1.02 \times 10^{-23} \mu M^{2/3} f^{2/3} r^{-1} \tag{8}$$

with a timescale (in seconds) of

$$\tau \approx 7.97 \mu^{-1} M^{-2/3} f^{-8/3}. \tag{9}$$

In the above equations, we see the total mass M and the reduced mass μ in units of the solar mass $M_\odot$, and the signal frequency f in units of 100 Hz [4]. Binary black hole (BBH) systems could therefore produce oscillating strain on the order of 10^{-21} for several seconds before merging, allowing their detection by gravitational-observatories like LIGO [13].

1.2 LIGO

The Laser Interferometer Gravitational-Wave Observatory (LIGO) Scientific Collaboration began running its latest generation of gravitational wave detectors in 2015. Referred to as the Advanced LIGO (aLIGO) detectors, these instruments are designed to achieve strain sensitivities on the scale of 10^{-23} around the 100 Hz frequency region, making it capable of observing binary black hole and binary neutron star coalescence, as well as other potential gravitational wave sources [3].

1.2.1 Laser Interferometry

The aLIGO detectors, located in Hanford, Washington and Livingston, Louisiana, are laser interferometers based on the Michelson interferometer, which makes use of light interference to measure distances. The basic design, as illustrated in Figure 2, is such that a laser beam is separated into two perpendicular arms by a beam splitter and then reflected by mirrors at the end of each arm [3]. The reflected light from both arms meets again at the beam splitter, where their amplitudes combine following the superposition principle and the resulting beam is sent to a photodetector. As the difference in the length of the two paths changes, an interference pattern in the resulting beam becomes apparent, allowing the measurement of the change in length to a precision limited only by quantum noise in the ideal limit [3].

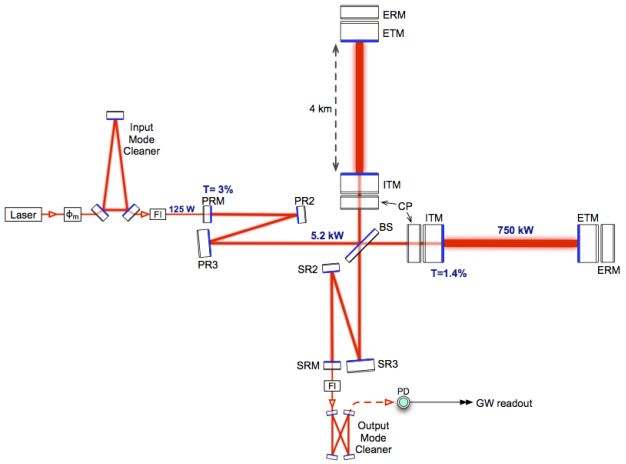


Figure 2: This schematic of the advanced LIGO interferometer shows the main elements of the optical path. The light leaving the laser is first passed through a phase modulator (ϕ_m), an input mode cleaning path, and a Faraday isolator (FI) which prevents interference in the input chain. The three power recycling mirrors, labeled PRM, PR2, and PR3, increase the effective power of the beam before it reaches the beam splitter (BS). Here the separated beams travel through the compensation plate (CP) and initial test mass (ITM) at the beginning of each "arm" of the interferometer. Upon reaching the end test mass (ETM), which is stabilized in part by the end reaction mass (ERM), the light is reflected back towards the beam splitter, where it interferes with itself and is sent to the signal recycling mirrors (SRM, SR2, and SR3). The final beam is then sent through a Faraday isolator and an output mode cleaning path, again to prevent unwanted interference, and is then converted into an electronic signal by a photodetector, labeled PD. Image credit: LIGO Scientific Collaboration [3].

The aLIGO interferometers differ from the Michelson interferometer by the addition of several resonant cavities and specialized components, thus optimizing the system to detect gravitational waves. Immediately noticeable is the immense length of the arms: 4 kilometers [3]. Such a long path is useful given that the aim is to measure the relative change in length of the arms. As the designed path length increases, the same precision in absolute length yields a greater precision in strain. There are also two important features that are not apparent in the schematic. The first is that the entire optical path is kept in high-vacuum, which reduces density fluctuations and scattering in the laser beam and promotes a constant effective refractive index along the path. The second is that each in-vacuum interferometer element is suspended by intricate vibration isolation systems to be detailed in section 1.2.3 [3].

1.2.2 Noise Sources

While in theory the precision of the interferometers is only limited by quantum noise (statistical fluctuations in photon number and detection rate), in practice there is a significant amount of noise from thermal vibration, residual gas, gravity gradients, and seismic activity. As we can see in Figure 4, the LIGO Scientific Collaboration has suppressed these noise sources such that in the 10 Hz to 7000 Hz detection band, quantum noise is the dominant contributor [3]. Thermal noise, caused by energy dissipation in the mirror coatings and test mass suspension fibers, is mitigated by using materials with low thermal response. The noise caused by residual gas particles in the beam bath is mitigated by operating in high-vacuum, as mentioned previously, such that the the only significant contributor is hydrogen with a pressure of 4×10^{-12} atm [3].

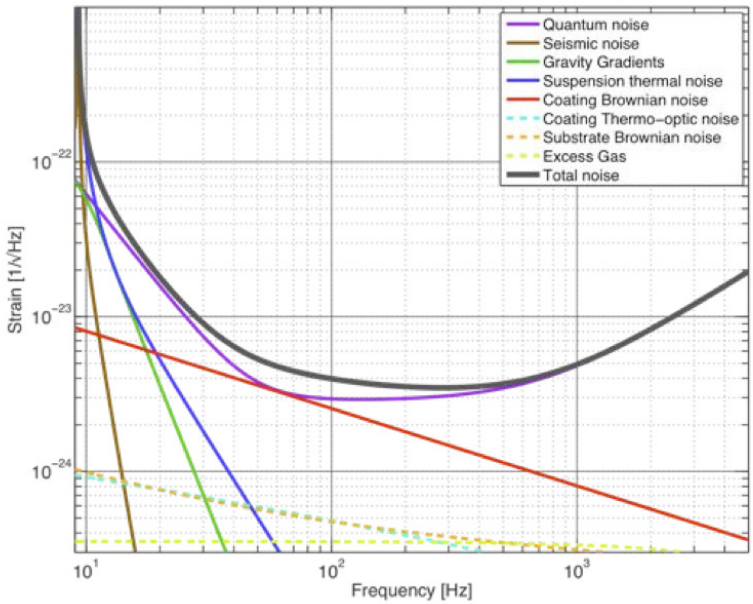


Figure 3: This image shows the different components that contribute to the estimation of the ideal sensitivity of the advanced LIGO detectors. It is apparent that the most significant source of noise is quantum "shot" noise, which arises from statistical inconsistency in the generation and subsequent detection of photons in the beam [3]. The strain noise level is measured in units of $1/\sqrt{\text{Hz}}$, as the figure is essentially displaying an amplitude spectral density of the theoretical detection level sans gravitational waves. Image credit: LIGO Scientific Collaboration [3].

1.2.3 Vibration Isolation Systems

The two-stage vibration isolation systems, designed to suppress noise from both seismic activity and local gravitational field gradients in all 6 degrees-of-freedom, are largely considered to be the most complex and challenging aspects of aLIGO's mechanical design. The first stage is an active isolation system which continuously monitors and corrects for ground vibration. Measurements of ground noise are taken with seismometers, geophones, and inductive position sensors located outside of the vacuum chamber, as well as seismometers and capacitive position sensors inside the chamber [3]. This allows two sub-stages of seismic isolation, one external to the chamber and another internal, using quiet hydraulic and electromagnetic actuators to achieve a strain noise level of 10^{-16} in the frequency region above 10 Hz [3].

The second stage of isolation is a passive multiple pendulum system, whose design varies slightly for each optical component of the interferometer. The most complex is the quadruple pendulum used to suspend the test masses, which mark the beginning and end of the interferometer arms [3]. In Figure 3, we see two adjacent chains of 4 masses with the test mass located at the bottom of the main chain. The top mass of each chain is actuated by the initial active isolation system, and a series of cantilevered blade springs and hanging masses passively rejects vibrations above 1 Hz, which brings the final strain precision to 10^{-22} in the most sensitive frequency band [3].

1.3 Motivation of Study

As we can see in Figure 4, the average noise contribution across all frequencies is expected to be significantly higher than the target sensitivity level, especially at very low frequencies. This necessitates complex signal processing and data filtering in order to successfully detect gravitational waves. Furthermore, much of the noise in aLIGO data is transient and non-Gaussian, so the analysis of potential signals must carried out case

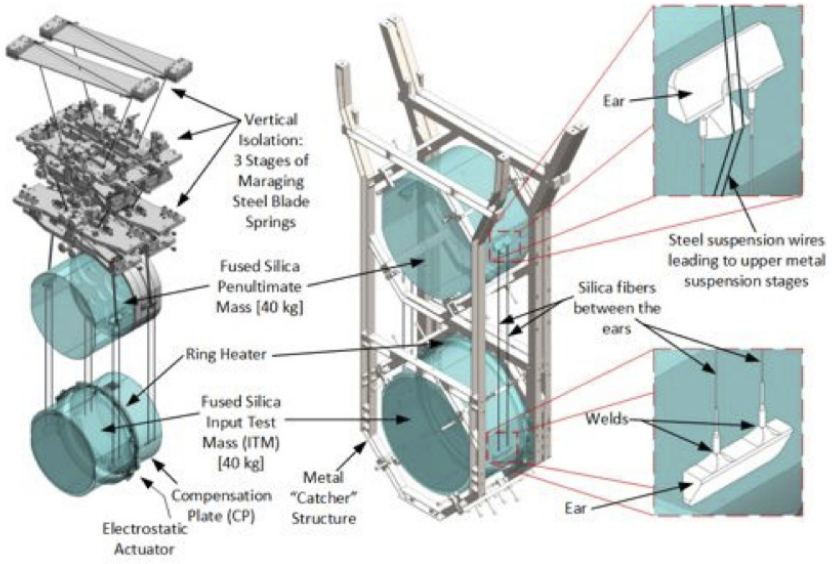


Figure 4: The above diagram shows the quadruple pendulum system which provides passive vibrational isolation for the input test mass. The top mass of the pendulum and the compensation plate are both subject to actuation, which further isolates the test masses from external vibration. Image credit: LIGO Scientific Collaboration [3].

by case, and even detector by detector [3]. Understanding the physical reasons for which various sections of the data can be neglected, as well as understanding the fundamental implications and assumptions that go into each filtering step, is imperative to the justification of the LIGO data analysis process as a whole. The remainder of this thesis is thus devoted to the exploration and clarification of each step in the signal processing "recipe" to search for gravitational waves.

2 Signal Processing

The LIGO Scientific Collaboration uses two independent search methods to identify candidate gravitational wave signals. The first involves flagging generic transient signals with high signal-to-noise ratio (SNR) by comparing time-shifted data between separate detectors. The second method employs matched filtering to compare incoming signals with a gravitational waveform template bank generated by analytical and numerical general relativity. Both of these methods rely on significant noise filtering in the frequency domain- namely whitening, bandpassing, and notch filtering- which is largely glossed over in recent literature [14].

2.1 Frequency Domain Filtering

The first step in the filtering process is to calculate the power spectral density (PSD) of the data in the time period just before and after the candidate signal. This gives us the best estimate of the background noise during the signal. It must be measured empirically as there is not currently a complete theoretical model of the detector and environmental noise, and it is impossible to completely isolate the detectors from gravitational waves and take pure noise measurements. Thus, the noise PSD is calculated by taking the time average of the square of the data "surrounding" the candidate signal, averaging over all detectors, and then extracting the frequency components via a Fourier transform [14]. In Figure 5 we can see that the amplitude spectrum density (ASD), which is the square root of the PSD, closely resembles the target noise estimations shown in Figure 4. The major difference is the addition of instrumental spectral lines that

arise from vibrational modes of the suspension fibers, electric power grid harmonics, and photon calibration methods.

Given the ASD, we can whiten the strain data of the candidate event in each detector by transforming to the frequency domain, dividing by the strain by the ASD, and then transforming back to the time domain. This effectively flattens the components of the signal which are part of the background noise, allowing low-amplitude signals to be detected amongst high-amplitude noise [14].

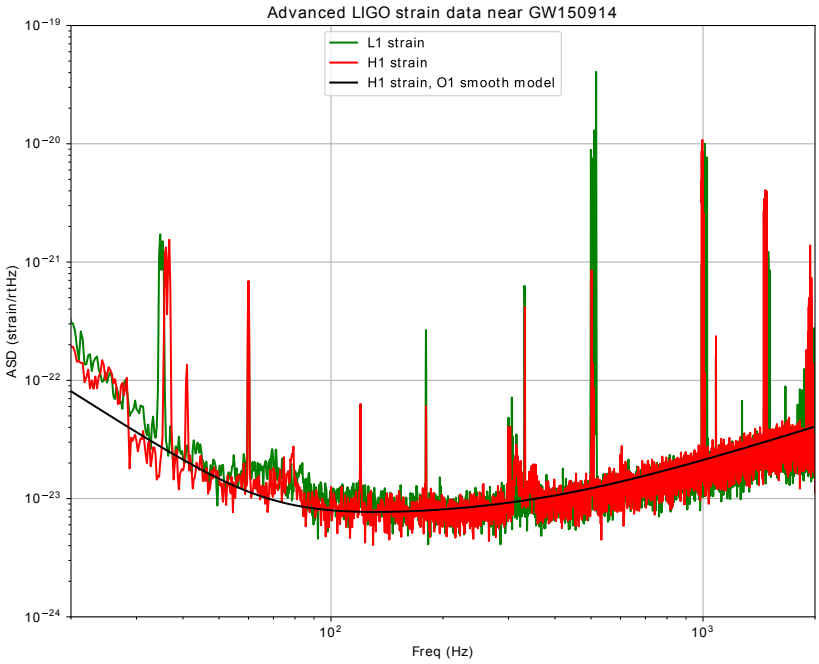


Figure 5: Here we show the amplitude spectral density (ASD) of the strain data from the 30 seconds around gravitational wave event GW150914, observed by the LIGO detectors in Hanford (H1) and in Livingston (L1). It clearly follows the same baseline shape as the predicted design sensitivity. The fundamental vibration mode of the quadruple pendulum appears as a strong spectral line at 510 Hz, with higher modes apparent at 1020 Hz, 2040 Hz, and so on. The features at 33-38 Hz and 330 Hz are due to photon calibration, in which secondary photon beams are reflected by the end test mass, creating predetermined displacements of the mirror [3]. The 60 Hz line and its integer multiples are due to electric power grid harmonics [1].

We can then use our prediction of the compact binary coalescence waveform shape to restrict the frequencies to a band spanning 10 Hz to 1 kHz. To do this we use a bandpass filter in the frequency domain, which rejects all frequencies outside of the allowed frequency range. The bandwidth of the filter is optimized case by case, as the frequency structure of the signal is highly dependent on the masses and spins of the merging objects [14]. Residual noise spikes in the frequency domain signal can be removed with notch filters, which reject only a narrow band of frequencies. These last two steps create a "smoothed" strain signal which can be compared between detectors.

The comparison step is invaluable because it allows further separation of the real strain signal from the noise background. The Livingston and Hanford detectors are separated by a physical distance equivalent to 10 ms of light travel time [1]. Therefore, identical gravitational wave signals that are apparent in both detectors within 15 ms of each other (with the extra 5 ms to account for uncertainty in the arrival time) are considered coincident events, and the SNR $\hat{\rho}$ for the combined event becomes

$$\hat{\rho} = \sqrt{\hat{\rho}_L^2 + \hat{\rho}_H^2}, \tag{10}$$

where $\hat{\rho}_L$ is the SNR of the signal detected in Livingston and $\hat{\rho}_H$ is the corresponding value in Hanford [14]. This equation can easily be expanded to include other gravitational wave detectors, such as the newly operational VIRGO detector.

2.2 Matched Filtering

The combination of frequency domain noise filtering and coincidence tests between detectors is enough to detect loud, clean gravitational wave signals. However, we can use the technique of matched filtering to detect weaker signals that match predicted gravitational waveforms emitted by compact binary coalescence (CBC) [15]. To achieve this, the signal is multiplied by a template waveform containing the target signal shape, and the result is integrated or summed. The final result, a scalar value, will be low when only noise is present, and it will be significantly higher when the target signal is present- even if the noise is louder than the signal itself. However, there are generally a number of unknown parameters that affect the final template waveform shape, and each of the different possible templates must be tested at every time-step in the data set. This results in a massive computational load and very long analysis times [15].

The use of Fast Fourier Transforms can simplify the process by allowing the matching process to run in the frequency domain as opposed to the time domain [16]. It has two main benefits: First, the time-series data can be processed for discrete chunks of time, rather than at each individual time-step, and second, the frequency analysis prevents high amplitude glitch signals from passing a high result as it would in the basic matched filter. After the Fast Fourier Transform of both the data and the template, each series is separated into narrow, mutually-exclusive frequency bands and then match filtered in the frequency domain. This ensures a high return only if the data matches the template at each frequency range, and not if the data simply has a high amplitude peak of arbitrary frequency /citematched.

2.2.1 Gravitational Waveform Template Generation

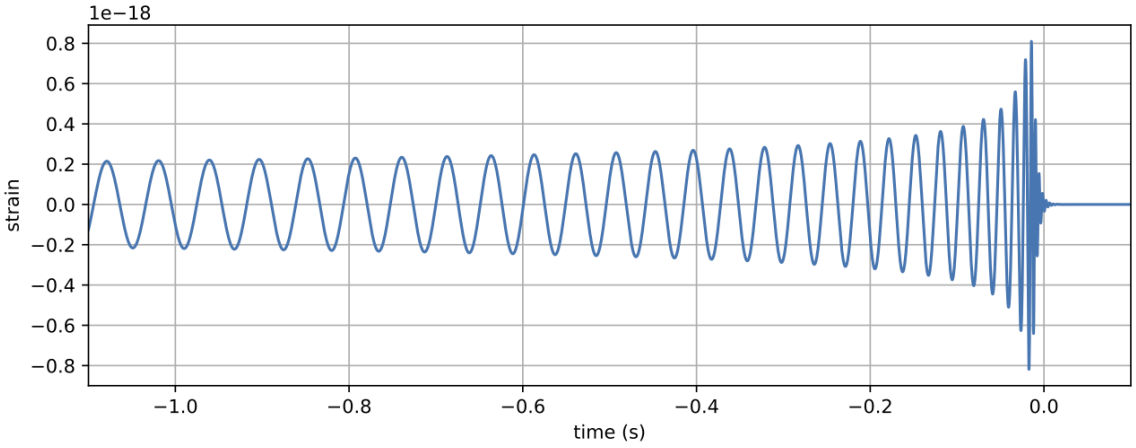


Figure 6: This plot shows a gravitational waveform template generated for the GW150914 event. The trend of rapidly increasing frequency and amplitude just before the merger is clear, followed by a brief ringdown period as the final object stabilizes [4]. Thousands of similar templates are generated to cover the relevant parameter space for compact binary coalescence.

The template used for a CBC search is a waveform generated by a combination of the post-Newtonian approach, black hole perturbation theory, and numerical general relativity concerning compact binary coalescence [16, 17]. It has a 4-dimensional parameter space composed of the individual masses and aligned spins of the two compact objects in the system. In constructing the templates, individual masses from 1 to 99 $M_\odot$ with a total mass under 100 $M_\odot$ and dimensionless spins up to 0.99 are considered, resulting in approximately 250,000 distinct templates.

The CBC waveform can be described qualitatively in multiple phases. The first is the *in-spiral* phase characterized by a quasi-sinusoidal signal with increasing frequency and amplitude. This corresponds to the period before the merger during which the compact objects are orbiting each other at a very close distance and with very high acceleration. The second phase is the *merger*. When the objects merge, the signal reaches a peak amplitude and frequency (called the merger frequency). The final phase, the *ringdown* phase is then characterized by rapidly decreasing amplitude at frequencies close to the merger frequency. This phase describes the final body reaching its equilibrium state [1]. An example is shown in Figure 6.

2.3 PyCBC and LALSuite

To accomplish the analysis described above, as well as many other processes, the LIGO Scientific Collaboration has developed a software framework called PyCBC. It is essentially a collection of Python wrapper functions designed to work with gravitational wave data. These wrapper functions call code from LALSuite, which is a collection of LIGO-designed object classes and data analysis routines written in C [14].

3 Results

The complete script to read raw strain data, flag potential gravitational wave detections, and analyze interesting transient signals is referred to as a pipeline. The LIGO Open Science Center (LOSC) provides a thorough pipeline to analyze simplified gravitational wave data using common Python packages like SciPy and NumPy [14]. However, the script must be rewritten using the LIGO-developed PyCBC software to handle the raw data from the aLIGO detectors. The following sections detail the construction of the optimal PyCBC pipeline and the effect of variations in the bandpass filtering steps.

3.1 Optimized PyCBC Pipeline

The optimal PyCBC pipeline was adapted directly from the LOSC Python script, as well as from the description in LIGO's first paper announcing the detection of GW150914. We first employ the PyCBC function "read frame," which reads the .gwf raw data file into a time series object. This object is a type of array that retains information concerning the sample rate and absolute time measurements, and it has a number of convenient methods such as "to frequencyseries," which applies a pre-constructed Fast Fourier Transform (FFT) to the time series. A plot of the raw data is shown in Figure 7, and we can see that no signal is immediately visible [14].

Here is where the matched filter is extremely useful. The applied code is as follows:

```
1   template = get_fd_waveform(approximant="IMRPhenomD" \
2      mass1 = 40, mass2 = 32, f_lower = 20, delta_f = 1.0/32)
3   template.resize(len(h1) / 2 + 1)
```

In this script, line 1 calls the "get fd waveform" function, which generates a gravitational waveform in the frequency domain using the approximant (approximate solution) type "IMRPhenomD" and the given mass, low frequency limit, and frequency step parameters in line 2. This template is then resized to match the length and resolution of the Hanford raw strain time series "h1" [14].

We then generate a PSD for the raw strain time series using the line:

```
4   psd = interpolate(welch(h1), 1.0 / 32)
```

This applies Welch's method to estimate the PSD of the raw time series "h1", and and then adjusts the resulting frequency series to have a frequency resolution of 1/32 Hz. The square root of the PSD generated here can be seen in Figure 5 for both the Hanford and Livingston data [14].

The matched filter itself is then applied with the following lines:

```
5   snr = matched_filter(template, h1, psd=psd, low_frequency_cutoff=20.0)
6   snr = snr[len(snr) / 4: len(snr) * 3 / 4]
```

In the above code, line 5 applies the "matched filter" function to the time series "h1," comparing it to the generated template at each time step. The PSD is used to weight the noise components, and the low

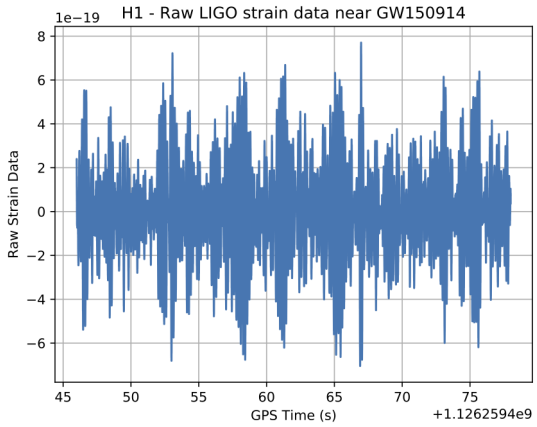


Figure 7: Here we display the raw strain data from the advanced LIGO detector at Hanford. No signal is immediately apparent.

frequency cutoff is set to 20 Hz, as the noise below this limit is very high. The matched filter returns a time series that measures the SNR as the template is matched with the data. The first and last quarter of the SNR time series is removed in line 6, as these regions are corrupted by the filter trying to match data outside the range of the H1 time series. The result of this step is shown in Figure 8.

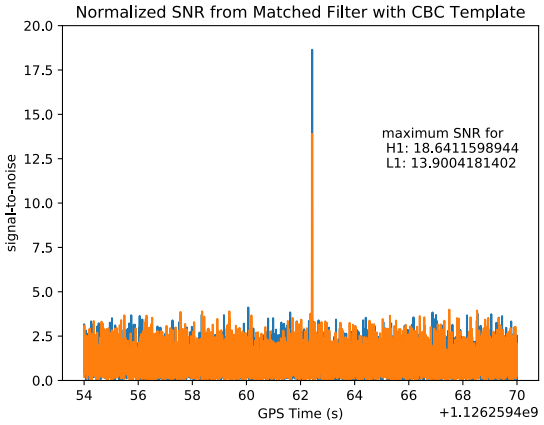


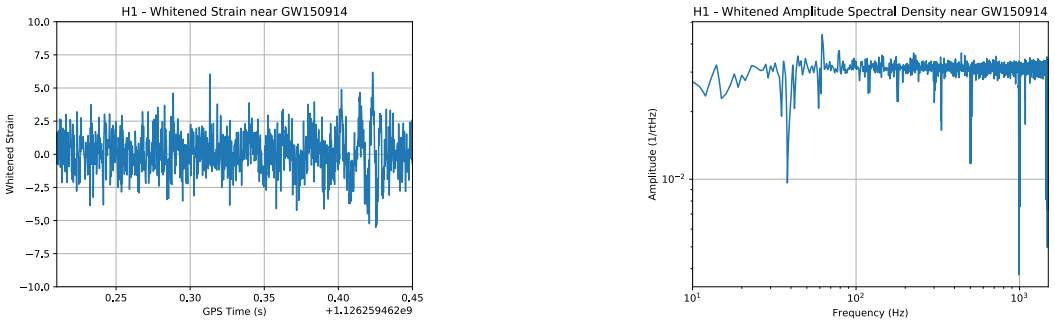
Figure 8: Here we display the result of the matched filter of both the Hanford and Livingston detections of the event GW150914, with a clear maximum denoting the presence of a CBC signal at 09:50:45 UTC on September 14, 2015. Due to the coincident detections, the total SNR for the GW event is 23.25 as given by equation 10.

The search for generic transient signals in the raw strain data involves a more complicated pipeline, because the data must be whitened, bandpassed, and in some cases notch filtered in order for any meaningful analysis to occur. The whitening step is achieved by the following code:

```
7    asd = psd ** 0.5
8    white_strain = (h1.to_frequencyseries() \
         / asd * asd.delta_f).to_timeseries()
```

In line 7, the ASD is calculated by taking the root of the PSD we found in line 4. Then, the whitened strain is calculated in line 8. We first apply the "to frequencyseries" method on the raw strain time series,

which transforms the time series to a frequency series using a FFT. The frequency series is then divided by the ASD and multiplied by the ASD frequency step to ensure that the result is normalized, and the result of this is transformed back to the time domain using the "to timeseries" method [14]. In cases where the gravitational wave is clear and the background vibration is relatively quiet, the signal becomes visible after this first step, as we can see in Figure 9.



(a) Time domain strain data

(b) Frequency domain strain data

Figure 9: These plots show the Hanford whitened strain data in both the time domain and the frequency domain, normalized against the amplitude spectral density. We can see that a quasi-sinusoidal signal is already visible in the time series data. We can also see the effect of whitening on the ASD: The regions of higher overall noise are flattened, and the weaker spectral lines are effectively notch filtered.

The next step is to bandpass the data, rejecting frequencies above and below the region of highest sensitivity, without removing too much of the signal. This is accomplished with the highpass and lowpass finite impulse response (FIR) filters that are built into PyCBC, shown here:

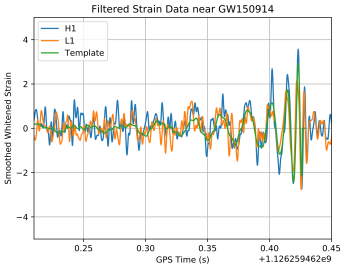
```
9    smooth_h1 = highpass_fir(white_strain, 35, 8)
10   smooth_h1 = lowpass_fir(smooth_h1, 300, 8)
```

This code applies a highpass and lowpass FIR filter of 8th order to the whitened strain to suppress frequency components outside of the range of 35-300 Hz. This is the bandpass used in both the LOSC python script and the LIGO Scientific Collaboration paper describing the detection of GW150914. The final waveform can be plotted in both the time domain and the frequency domain to check for residual noise spikes which require additional filtering. In the event that the signal is clear in both the Hanford and Livingston detectors, then the gravitational waveform template used in the matched filter can undergo the same whitening and bandpassing process to be used as a fit for the final waveform [14]. The results of the optimized waveform are shown in Figure 10.

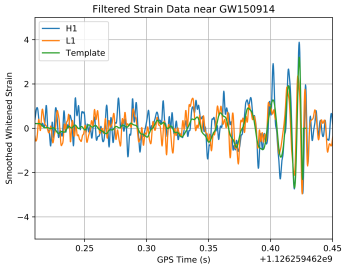
(a) Time domain strain data

(b) Frequency domain strain data

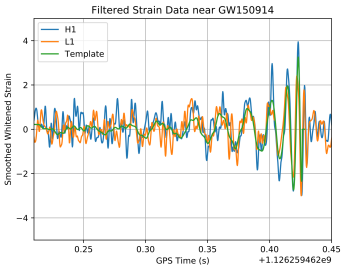
Figure 10: The above plots show the whitened strain data near the gravitational wave event GW150914, smoothed with a 35-300 Hz bandpass filter. The overlaid template in plot (a) has been filtered identically, and determines the merger frequency to be approximately 250 Hz.



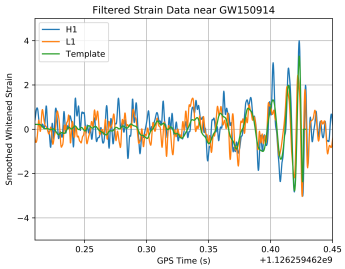
(a) Bandpass 35 - 300 Hz



(b) Bandpass 10 - 300 Hz



(c) Bandpass 5 - 300 Hz

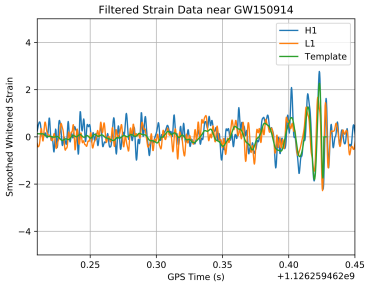


(d) Bandpass 1 - 300 Hz

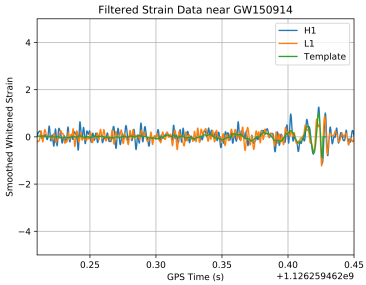
Figure 11: The above plots show the whitened and filtered strain data as the low frequency limit is decreased from 35 Hz to 1 Hz. This seems only to have the effect of a slight increase in the strain amplitude. The signal to noise ratio remains nearly the same.

3.2 Effect of Filter Variations

Given that the bandpassing steps are performed case-by-case, we explore the effect of varying the upper and lower frequency limits imposed on the strain data. The effect of can be seen in the final waveforms as the amplitude of the CBC signal changes. As the lower frequency limit is decreased from 35 Hz to 10 Hz, and then to 1 Hz, one would assume that the ground noise would begin to overpower the signal. However, we can see in Figure 11 that the signal remains nearly unchanged. This leads us to understand even as low as 1 Hz, the combination of the vibration isolation system and the whitening filter is sufficient to suppress external noise sources.



(a) Bandpass 100 - 300 Hz

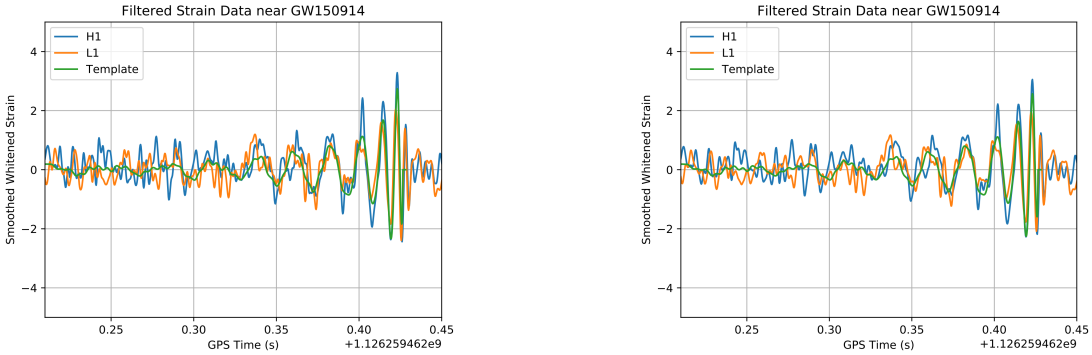


(b) Bandpass 250 - 300 Hz

Figure 12: The above plots show the whitened and bandpass filtered strain data as the low frequency cutoff is increased to 100 Hz and to 250 Hz. This significantly decreases both the amplitude and SNR of the CBC signal.

While the strain amplitude becomes larger as the low limit approaches 1 Hz, the frequency components contributing to that increase are likely due to noise rather than gravitational wave strain. LIGO has also

stated that its data below 10 Hz is not well calibrated, which further promotes a low frequency limit above 10 Hz. As we increase the low frequency cutoff to 100 Hz and to 250 Hz, we see in Figure 12 that the filter begins to remove portions of the gravitational wave signal.

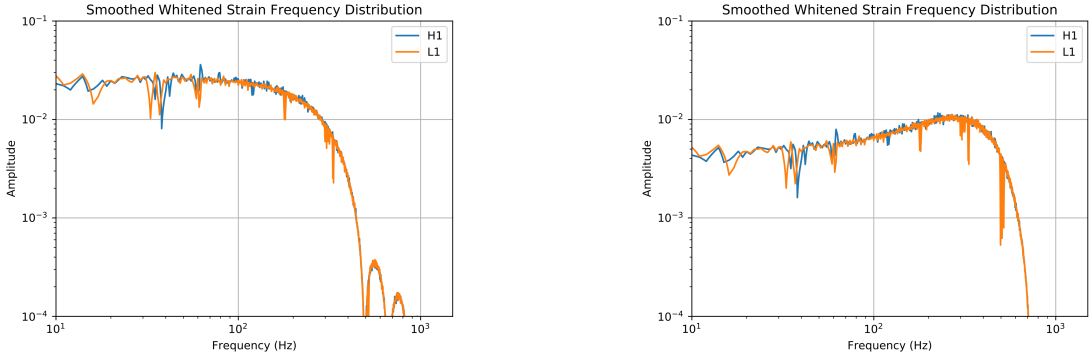


(a) Bandpass 35-200 Hz

(b) Bandpass 35-40 Hz

Figure 13: These figures show the whitened and bandpass filtered strain data as the high frequency cutoff is decreased to 200 Hz and to 40 Hz.

Upon decreasing the upper frequency cutoff of the bandpass filter, we find an analogous result. As the upper limit is decreased to 200 Hz, and to 40 Hz, the strain amplitude decreases. However, the amplitude after the 35-40 Hz bandpass is nearly three times that of the 250-300 Hz bandpass, even though a much smaller frequency range is accepted. This prompts the examination of the amplitude in the frequency domain of these two waveforms.

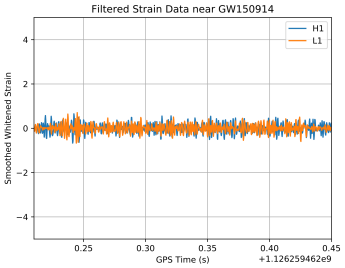


(a) Bandpass 35-40 Hz

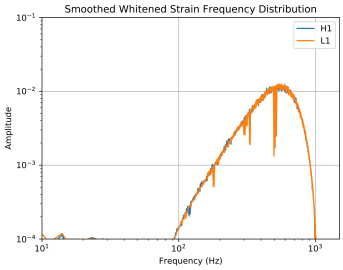
(b) Bandpass 250-300 Hz

Figure 14: Here we show the amplitude spectral density of the strain data after a 35-40 Hz bandpass and a 250-300 Hz bandpass. We can see that the amplitude in the frequency range outside the bandpass is not suppressed very quickly, especially at low frequencies.

Figure 14 reveals that the high- and low-pass FIR filters do not provide sharp cutoffs for frequency content. They provide instead a suppression whose strength increases as the frequencies become more distant from the cutoff value. We can see that in the case of the 300 Hz upper limit, the content at 500 Hz is only decreased approximately half of the maximum amplitude, and it decreases to one tenth at 600 Hz. And the 250 Hz lower limit still allows approximately half of the maximum amplitude to pass through as low as 10 Hz. If we look at the result of a 500-600 Hz bandpass, which completely removes the CBC signal as depicted in Figure 15 (a), we can see in the frequency domain a consistent "bleed" of approximately 300 Hz beyond both the low and high limit, where the "bleed" is defined by admission greater than 10%. This explains the presence of a loud signal even when we have only defined a 5 Hz gap to be admitted.



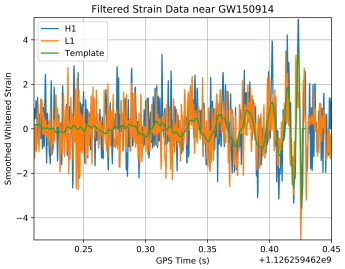
(a) Time domain strain data



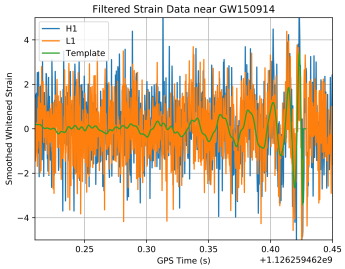
(b) Frequency domain strain data

Figure 15: The above figures show the time domain and frequency domain amplitudes of the 500-600 Hz bandpass result. We can see the response of the whitened frequency spectrum to the bandpass filter in a region where there is no signal. This clearly shows the the additional 300 Hz range extending from each end of the desired frequency range for which there is greater than 10% admittance.

We then observe the result of increasing the upper frequency limit. As depicted in Figure 16, high cutoffs of 1 kHz and 2 kHz result in a significant increase in background noise and a significant decrease in SNR. This is to be expected, as we know the maximum frequency present in the CBC signal is the merger frequency at approximately 250 Hz.



(a) Bandpass 35-1000 Hz



(b) Bandpass 35-2000 Hz

Figure 16: This plot shows the whitened and filtered strain data as the upper frequency cutoff is increased to 1000 Hz and to 2000 Hz. We can see that a significant amount of noise is admitted, increasing the strain amplitude and decreasing the SNR.

4 Discussion

The exploration of the PyCBC pipeline in the above sections has served to illustrate two main points. The first is that we are able to determine the exact effects and implications that the various filtering steps impose on CBC strain data, and that we can explain why these effects are either negligible or inconsequential to the strength of the GW signal. The second is that we are able to define limits for the bandpass ranges empirically, that is, without relying on a theoretical model. This is especially important in the search for generic transient signals, as we would not be able to use matched filtering to seek out predetermined waveforms.

4.1 Rationalization of Filtering Steps

The first step in both the matched filtering search and the generic transient search (which is effectively matched filtering between multiple detectors) is the noise weighting or whitening step. We saw in Figure 9 (a) that this step is already a thorough enough filter to render strong signals visible in the time series. We also saw in Figure 9 (b) that the whitening step effectively flattens regions of high overall noise and

applies notch filters that account for spectral lines in the noise ASD. Because the strength of the whitening is directly correlated to the amplitude of the noise present in the detector, the algorithm does not need to be adjusted for particular waveforms or types of transient signals.

The bandpass filters must be applied with more care, as their effects depend entirely on user input rather than measured parameters. In the case of a CBC search, the gravitational waveform template that triggers the matched filter already contains information about the frequency content of the signal, and the range of frequencies present in the template can be used as a logical starting point for the bandpass filter. The template used to match the event GW150914 is shown in Figure 17. This template ASD's frequency cutoff at 250 Hz reveals that signal content above that frequency is likely to be noise, and that removing it will not weaken the strength of the GW signal.

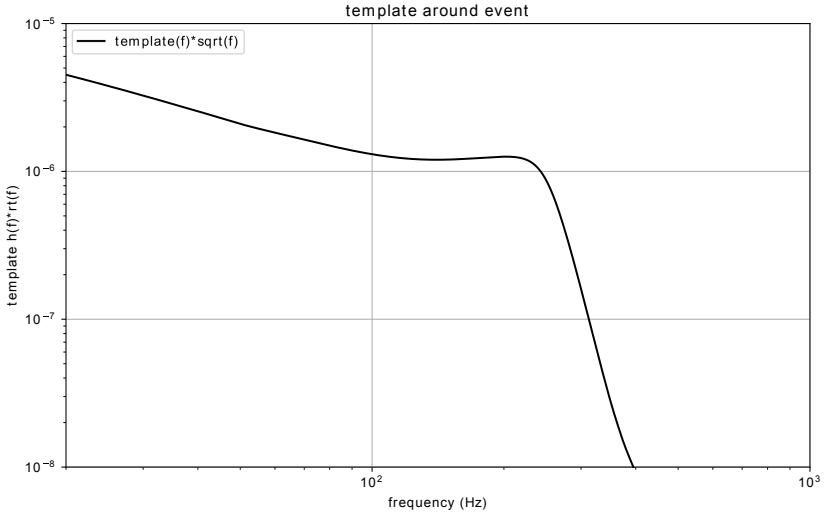


Figure 17: This plot shows the amplitude spectral density of a gravitational waveform template generated for the GW150914 event. We can see evidence of a maximum frequency present in the signal at around 250 Hz, as well as a high signal content at low frequencies. Therefore it is reasonable to use a high frequency cutoff at 300 Hz (which will not remove any significant signal) and a low frequency cutoff of 35 Hz which will mediate low frequency noise while only affecting the earlier (lower frequency) part of the inspiral signal, depicted in Figure 6.

4.2 Limits on Bandpass Filters

In the case of a generic transient search, after a coincident trigger has been confirmed between detectors, there are no exact guidelines concerning what frequency range the potential signal is present in. An array of different bandpass filters may then be applied to find the one with the highest resulting SNR, provided the filters used have a solid physical motivation. The low frequency cutoff can be motivated by the presence of strong ground noise or low frequency "blip glitches," which are transient noise spikes of unknown origin in LIGO detectors [18], as well as increasing quantum noise at higher frequencies.

When working with LIGO data, a sensible start would be to apply a bandpass admits frequencies within a generous estimate of detector's most sensitive range: on the order of 10 to 1000 Hz. If we had begun the analysis of the GW150914 event with this method, the result of this filter would look like figure 18. In comparison with the whitened strain data shown in Figure 9 (a), the bandpass does not lend much more clarity to the data, but neither does it obscure the data. We know now that the FIR bandpass filters leave "soft" edges in the ASD's of the filtered data, and that during times of quiet background, the whitening step largely takes care of low frequency noise. We can see from the results of this experiment, and from the discrepancy between the choice of bandpass made in LIGO's published scientific papers (35-350 Hz) and in

the LOSC tutorial software (43-300 Hz), that the frequency range chosen for the filter does not have to be exactly precise in order to locate a statistically significant signal [1, 4].

(a) Time domain strain data

(b) Frequency domain strain data

Figure 18: The above figures show the time domain and frequency domain amplitudes of the 10-1000 Hz bandpass result.

The combination of matched filtering, which makes comparisons among signals in the frequency domain, with bandpass filtering, which suppresses signals in unwanted frequency ranges, gives us incredible insight into the seemingly empty signal that we showed in Figure 7.

5 Conclusion

There exists no real signal detected by a real instrument that is completely free of noise. While background noise and uncertainty can be negligible in some situations, in the field of gravitational wave astronomy, the amplitude of quantum, seismic, and thermal vibrations are often comparable to or greater than the target signal representing spatial strain. To measure the distortion of space itself, researchers at the LIGO Scientific Collaboration (LSC) have employed advanced data processing and filtering techniques.

The completion of this thesis has resulted in a clearer understanding of how signals, both known and unknown, can be identified in spite of significant background noise. The techniques of matched filtering and bandpass filtering make use of the distinct frequency makeup of gravitational waves, and by using these techniques in combination with highly advanced interferometry, vibration isolation, and other technologies, the LSC has made the first ever direct detection of the spatial strain caused by gravitational waves. The observed event, GW150914, had a peak strain amplitude on the order of 10^{-21}, and its detection represents the culmination of nearly 100 years of scientific advancement, and confirms predictions made by Albert Einstein's theory of general relativity, published in 1916. As gravitational radiation is independent of electromagnetic radiation, it offers us an entirely new source of information about the universe, and especially about dark matter, which does not interact with the electromagnetic force. The nature of gravitational wave astronomy lends itself to inspiring ever more advanced precision technology and signal processing techniques, and its potential for new discovery is boundless.

References

[1] B. P. Abbott *et al.*, "Observation of gravitational waves from a binary black hole merger," *Phys. Rev. Lett.*, vol. 116, p. 061102, Feb 2016.

[2] P. R. Saulson, "Josh goldberg and the physical reality of gravitational waves," *General Relativity and Gravitation*, vol. 43, pp. 3289–3299, Dec 2011.

[3] T. S. C. LIGO *et al.*, "Advanced ligo," *Classical and Quantum Gravity*, vol. 32, no. 7, p. 074001, 2015.

[4] S. LIGO, C. VIRGO, *et al.*, "The basic physics of the binary black hole merger gw150914," *Annalen der Physik*, vol. 529, no. 1-2, 2017.

[5] C. Misner, K. Thorne, and J. Wheeler, *Gravitation*. W. H. Freeman, 1973.

[6] C. D. Hill and P. Nurowski, "How the green light was given for gravitational wave search," *Notices of the American Mathematical Society*, vol. 64, no. 7, pp. 686–692.

[7] A. Einstein, "Die grundlage der allgemainen relativitätstheorie," *Annalen der Physik*, vol. 49, no. 7, pp. 769–822.

[8] S. S. Chandrasekhar, *The mathematical theory of black holes*. Oxford classic texts in the physical sciences, 1998.

[9] A. Einstein, "Über gravitationswellen," *Sitzungsberichte*, p. 154–167, 1918.

[10] I. Robinson and A. Trautman, "Spherical gravitational waves," *Physical Review Letters*, vol. 4, no. 8, pp. 431–432.

[11] H. Bondi, "Plane gravitational waves in general relativity," *Nature*, vol. 179, no. 4569, pp. 1072–1073.

[12] B. F. Schutz *Nature*, vol. 310, 1986.

[13] F. Pretorius, "Evolution of binary black-hole spacetimes," *Phys. Rev. Lett.*, vol. 95, p. 121101, Sep 2005.

[14] S. A. Usman *et al.*, "The pycbc search for gravitational waves from compact binary coalescence," *Classical and Quantum Gravity*, vol. 33, no. 21, p. 215004, 2016.

[15] B. J. Owen and B. S. Sathyaprakash, "Matched filtering of gravitational waves from inspiraling compact binaries: Computational cost and template placement," *Phys. Rev. D*, vol. 60, p. 022002, Jun 1999.

[16] B. F. Schutz, "Gravitational wave astronomy," *Classical and Quantum Gravity*, vol. 16, no. 12A, p. A131, 1999.

[17] A. Taracchini *et al.*, "Effective-one-body model for black-hole binaries with generic mass ratios and spins," *Physical Review D*, vol. 89.

[18] A. H. Nitz, "Distinguishing short duration noise transients in ligo data to improve the pycbc search for gravitational waves from high mass binary black hole mergers," *Max-Planck-Institut für Gravitationsphysik, Albert-Einstein-Institut*, September 2017.

[19] S. Carroll, "A no-nonsense introduction to general relativity," *University of Chicago*, 2001.

[20] J. P. A. Clark and D. M. Eardley, "Evolution of close neutron star binaries," *The Astrophysical Journal*, vol. 215, pp. 311–322.